DOCUMENT RESUME

ED 078 672 EM 011 216

AUTHOR Avner, R. A.
TITLE Objective Criteria for Evaluation of Grading Scales.
INSTITUTION Illinois Univ., Urbana. Computer-Based Education Lab.
SPONS AGENCY National Science Foundation, Washington, D.C.; Office of Naval Research, Washington, D.C.
REPORT NO CERL-R-X-18
PUB DATE Aug 70
NOTE 20p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Achievement Rating; Classification; *Evaluation; *Evaluation Criteria; *Grading; Measurement; Measurement Goals; *Norms; Pass Fail Grading; Predictive Measurement; Statistical Analysis

ABSTRACT

This report compares maximum linear prediction, maximum total correct classifications for a group, and maximum probability of correct classification for an individual as objective criteria for univariate grading scales. Since the goals of valid prediction and valid classification lead to conflicting criteria, it is possible that a compromise measure consisting of a standardized score with a variable pass-fail point may provide the best measure available. (Author/PB)

CERL REPORT X-18 AUGUST, 1970

OBJECTIVE CRITERIA FOR EVALUATION OF GRADING SCALES

R. A. AVNER

Computer-based Education Research Laboratory

University of Illinois Urbana Illinois

ED 078672

This work was supported in part by the Advanced Research Projects Agency under grant ONR Nonr 3985(08) and in part by the Office of Education, U.S. Department of Health, Education, and Welfare under contract no. OE-6-10-184.

OBJECTIVE CRITERIA FOR EVALUATION OF GRADING SCALES

The wide use of sophisticated multivariate measures sometimes leads us to overlook possible margins for improvement in the component univariate measures. The academic grade is, from one standpoint, an example of such a univariate measure. A single grade is typically used to summarize a multivariate collection of measurements and judgments of student behavior in a given course. Frequently, there is some intermediate measure such as the numerical sum of scores given on various tests and evaluations. The objective here will be to review some ways to assess the loss in information contained in these intermediate univariate measures through transformation to sub-optimal univariate grade scales. The loss of information due to combination of multivariate measures into a univariate measure will not be covered. In the course of this review, explicit application will be made of some representative criteria which appear to be implicitly used by those who have selected certain presently used grade scales.

Criteria appear to be of two types; (1) those leading to grades which imply a continuum of performance and (2) those which imply discrete categories of performance. An example of criteria which imply a continuum of performance is the product-moment correlation coefficient as used by Willingham (1964) to measure effects of use of different grading scales on validity coefficients. Simple objective criteria of categorizations are more difficult to find but measures such as the total number of correct categorizations are representative.

Since some type of 5-category grading scale (e.g. A-B-C-D-E) is probably most widely used in U. S. colleges we will begin by evaluating a selection of such scales.

Category Size

One approach to the translation of test scores into grade categories is to define the grade category size in terms of standard deviation (s.d.) units of the observed score distribution. If the underlying distribution is normal it is obvious that only the middle three categories of a 5-category scale can conform to the specified size. Since the normal distribution is unbounded the two "outer" categories must each have one bound unspecified so as to include all extreme values. If we decrease the size of the three central categories we find that as the category size approaches zero, the central categories tend towards insignificant size and we obtain a 2-category scale (the extreme categories) as a limiting case. On the other hand, with increasing central category size (hereafter abbreviated "CCS") we find that more and more observations fall into the center category, a 1-category scale is ultimately approached. As can be seen in Table 1, by the time a CCS of about 1.6 s.d. is reached the grade scale has become effectively a 3-category scale.

TABLE 1

Distribution of Ss on 5-category scales

of Varying Central Category Sizes

	Category				
CCS	A	B	C	D	E
.0	.50	.00	.00	.00	.50
.2	.38	.08	.08	.08	.38
.4	.27	.15	.16	.15	.27
.6	.18	.20	.24	.20	.18
.8	.12	.23	.31	.23	.12
1.0	.07	.24	.38	.24	.07
1.2	.04	.24	.45	.24	.04
1.4	.02	.22	.52	.22	.02
1.6	.01	.20	.58	.20	.01

Note.--CCS is in standard deviation units.

An application of the criteria mentioned above would be the selection of one of the category sizes in this range as optimal for some related purpose.

Grade Scaling Model

By use of Pearson's tables of volumes under a bivariate normal surface for different values of ρ (Pearson, 1931, Tables VIII and IX), it is possible to construct 5 x 5 contingency tables which represent the number of joint categorizations on two scales having a population correlation of ρ and a given sample size N. Such tables were constructed for N=1,000,000, ρ = 0, .50, .70, .95, and 1.00 and for the nine CCS values used in Table 1. Table 2 shows one such table. In each case it is assumed that the scores on the two measures being compared are

TABLE 2

Sample Contingency Table

Joint Distribution of Scores on Two Scales

A	0	0	16	16237	50554
B	0	27	44512	180954	16237
C	16	44512	293870	44512	16
D	16237	180954	44512	27	0
E	50554	16237	16	0	0
	E	D	C	B	A

Note.—ρ=.95, N=1,000,000, CCS=1.0

based on an underlying continuous bivariate normal distribution with parameter ρ. In the following discussion examples using ρ = .95 will be emphasized since this is close to the value of .94 suggested by Kelly (1927) as the minimum degree of relationship necessary for measures of individual accomplishment.

Correlation

For each contingency table we can compute the product-moment correlation coefficient, r. Figure 1 shows the relationship of r to CCS

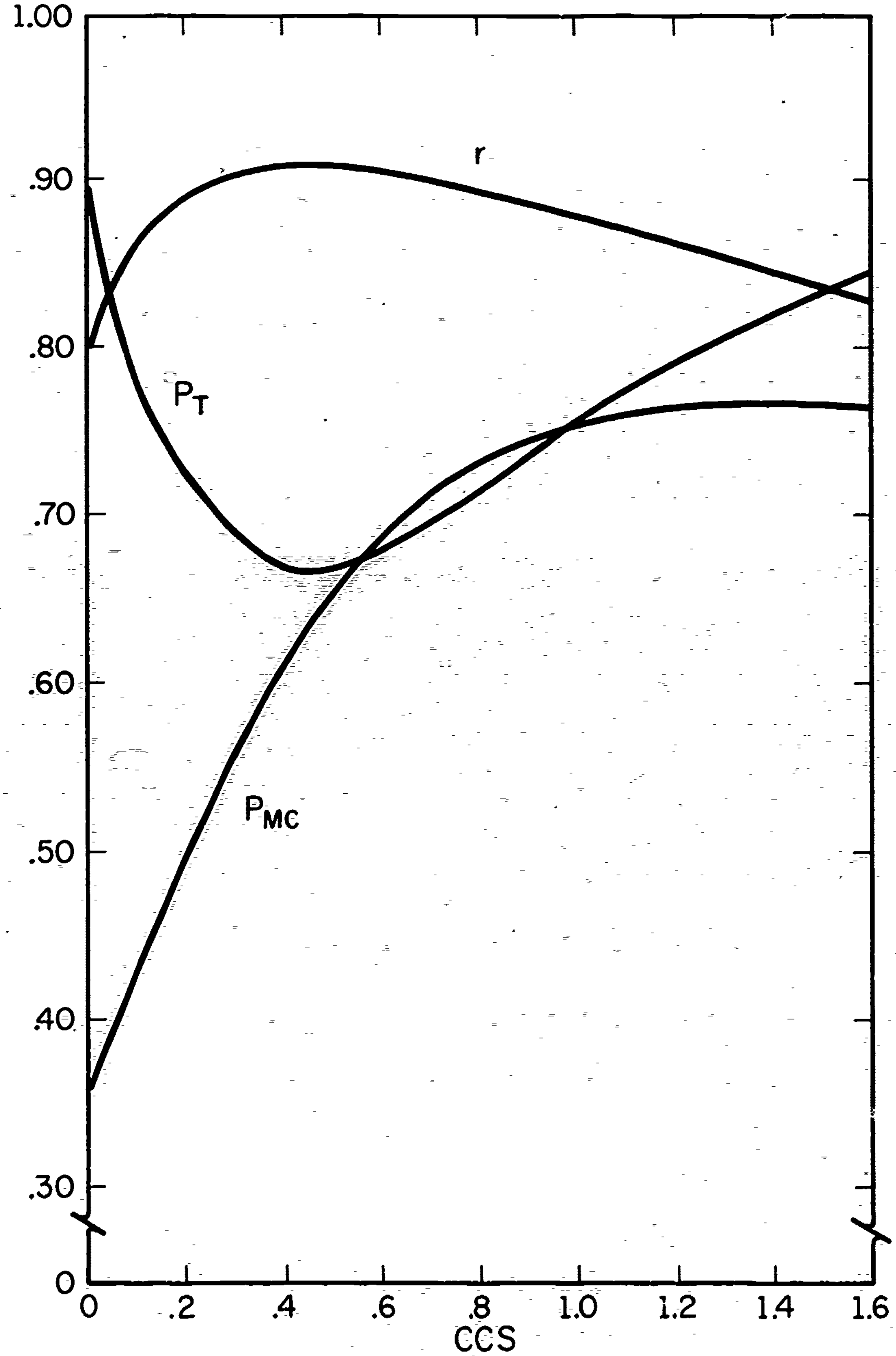

Figure 1. Evaluation of 5-category scales of varying Central Category Size by three criteria. Underlying population is bivariate normal with $\rho = .95$.

for $\rho = .95$. It will be noted that r is less than ρ for all values of CCS. There are "coarse grouping" formulas available for correcting r as an estimate of ρ (e.g. Peters & Van Vorrhis, 1940) but aside from the fact that they produce poor approximations as ρ approaches unity (Kelly, 1947), they are of little value here since our interests lie in the relationship exemplified by r rather than by ρ.

If r is to be maximized, Figure 1 indicates the optimum value of CCS (for $\rho = .95$) is about .4 s.d. For lower values of ρ the optimum value of CCS increases and at $\rho = .50$ it is about .8 s.d.

The actual interpretation of r of course depends on how the two scales represented by the contingency tables are defined. If the two scales are defined as equivalent imperfect measures, r is a reliability coefficient. If one scale is an imperfect predictor score (e.g. freshman grade) and the other scale is an imperfect criterion score (e.g senior grade), we have a validity coefficient, and so forth. In general the magnitude of r is directly proportional to the number of categories used. Hence, if grades with maximum reliability and validity are desired and the underlying distribution of performance is normal, grade scales with large numbers of categories must be used. As Ebel (1965) has pointed out, it is a fallacy to believe that one can reduce the effect of unreliable measures by reducing the number of grade categories. Reduction of the number of categories simply reduces the reliability of the measure still more.

Total Correct Classifications

The diagonal cells in Table 2 for which the same grade occurs on both scales contain all of the "true" classifications. The proportion of observations falling into these correct cells (P_T) is a second criterion for evaluating scales. Ebel (1947) has shown that although this criterion varies directly with ρ it is otherwise of questionable value. Figure 1 shows the relationship of P_T to CCS for $\rho = .95$. Comparison of curves for r and P_T show that maximum r and minimum P_T occur at almost the came CCS for a given ρ . Maxima occur for P_T when CCS is zero (the 2-category scale limiting condition) or infinite (the

1-category scale limiting condition). The latter maximum is unity for all ρ (i.e. it is impossible to misclassify when only one category is available) while the former is unity only for $\rho = 1$. In general the value of P_T (unlike r) is inversely related to the number of categories used. Hence, if the aim is to maximize the total number of correct classifications, one should use the minimum possible number of categories.

A major drawback to using maximum P_T as a criterion for evaluating grading scales can be seen in Table 3 which compares two 3-category (e.g.

TABLE 3

Classification Probabilities on Two 3-Category Scales

CCS	P_T	P(B)	$P(B_1 \mid B_2)$
.2	.84	.08	.25
1.0	.82	.38	.77

Note.--P(B) is the probability of receiving the middle grade (B) on either scale. $P(B_1 \mid B_2)$ is the conditional probability of receiving a B on both scales. $\rho = .95$.

A-B-C) scales. The scale with CCS = .2 s.d. is the better, based on larger P_T. However, because the central category, "B", is assigned with a probability of only .08, the individual with a true score of "B" is three times more likely to have an observed score of "A" or "C" than he is to have an observed score of "B". Thus the criterion of minimum total number of classification errors can lead to high probabilities of incorrect classification for individuals falling into central categories when CCS is small. On the other hand, when CCS is large, high P_T is gained by placing almost everyone in the middle category, thus retaining a minimum of the information available in the original measure.

Mean Conditional Probability

Table 4 is the validity matrix (Cronbach & Gleser, 1957) constructed from the information in Table 2. Table 4 is simply a table of the conditional probabilities of an observation falling in any category of one scale, given that it is in a specified category on the other scale (e.g. the probability of receiving an "A", "B", "C", "D", or "E" in course II given that a "B" is received in course I). The values of the cells on the major diagonal are thus the probabilities of receiving the same grade on both scales. Maximization of the expected value of the probabilities of receiving the same grade on both scales (P_{MC}) is a third possible criterion. Like $\underline{r}$ and P_T, P_{MC} increases when the value of ρ increases. P_{MC} ranges from unity when ρ is unity to $^1/_n$ (for an n-category scale) when ρ is zero. For a 5-category scale if $\rho = .95$, maximum P_{MC} occurs when the CCS is about 1.2 s.d. This optimum point can be seen in Figure 1 which compares P_{MC}, P_T and $\underline{r}$ for $\rho = .95$.

TABLE 4

Sample Validity Matrix

Distribution of Scores on Scale II for Each

Score on Scale I

	A	.0000	.0000	.0002	.2430	.7567
	B	.0000	.0001	.1841	.7486	.0672
I	C	.0000	.1162	.7674	.1162	.0000
	D	.0672	.7486	.1841	.0001	.0000
	E	.7567	.2430	.0002	.0000	.0000
		E	D	C	B	A
				II		

Note.-- $\rho = .95$, CCS = 1.0.

P_{MC}, unlike P_T is unaffected by weighting effects of differential category occupancy. Being a function of the total number of correct classifications, P_T is of greatest value as a criterion for group decisions while P_{MC}, giving equal weight to all categories, is of greatest value as a criterion for making individual decisions. Thus, to minimize the total number of misclassifications P_T should be used but to minimize the probability that any individual is misclassified, P_{MC} should be used. Unfortunately, P_T and P_{MC} lead to the same results only in the special case of uniform distribution of observations across all categories.

Comparison of Criteria

When an attempt is made to find the best 5-category scale under the assumption of an underlying bivariate normal distribution of grades, maximization of P_T leads to results which are probably unacceptable in practice. Maximization of either r or P_{MC} leads to distributions which are acceptable but not identical. It is interesting to note that a value of CCS commonly used for 5-category scales (1 s.d.) lies between the optimum value dictated by use of r and that dictated by use of P_{MC}.

Even at face value, neither of the two categorization criteria would be likely to be acceptable as the sole criterion for development of a scale. Categorization is generally done for some immediate purpose, e.g. to determine those who must repeat the course, those who may continue with the next course, and those who are so advanced that they can skip to beyond the next course. The categorization criteria given here, if followed alone, would provide reliability without guaranteeing validity. Therefore, rather than being used to dictate the proportion of Ss to be placed within each category, the categorization criteria are more useful in evaluating placements performed under more complex criteria.

An alternate approach in which the effect of varying the number of scale categories (n) while examining r and P_{MC} is shown in Table 5. It

TABLE 5

P_{MC} and r as a Function of n (number of scale categories)

n	1	2	3	4	5	7	9
P_{MC}	1.00	.90	.82	.79	.76	.71	--
r	--	.80	.89	.89	.91	.93	.93

Note.--CCS=1 for all P_{MC} and CCS=.4 for all r, ρ = .95.

is apparent from Table 5 that P_{MC} dictates the use of scales having a minimum n while r dictates the use of scales which have a maximum n. Since it is always possible to transform a scale into a new scale having fewer categories (e.g. A-B-C-D becomes "Pass" while E becomes "Fail"), a scale which is efficient for prediction can generally be reduced to an efficient scale for categorization. The reverse is not true. Hence, if either prediction or both prediction and categorization are required, it is desirable to have a scale with the largest practical n.

If only categorization is required it is obviously necessary to have a scale which has n at least as large as needed in the ultimate categorization operation. No harm accrues by having n larger than this minimum so long as the scale can be transformed prior to ultimate use.

Reference to Table 5 indicates that r increases very slowly for $n>7$ (For ρ = .95, CCS = .4, going from n = 7 to n = 9 increases r by only about .007). In this range of ρ the optimal CCS is about .4 s.d. while for lower values of ρ the optimal CCS will increase slightly. This suggests the stanine scale as a practical alternative to the present 5-category scale. In the terminology of this paper the stanine scale has n = 9 and CCS .5 s.d. Since the conversion of raw scores to stanine scores is also a normalizing transformation, the marginal distributions become normal and the assumptions of normality made up to this point will tend to be supported if the joint distribution is unimodal. The appendix to this paper shows the effect of using a stanine grading scale on some actual data.

Conclusions

Neither the 5-category nor the 3-category scale is optimal for either the prediction or the usual (pass-fail) categorization problem. Every single-score grading system which must provide information for both of these objectives is a compromise. The "pass-fail" marking systems reduce the task of the instructor t[illegible] critical decision rather than the four required for a 5-category system but at the same time act as a filter for much of the predictive information available from classroom measures. "Curving" grades to fit a normal distribution with a predetermined pass-fail point removes the requirement for any instructor's decision while providing for a fair amount of linear prediction at the cost of possibly invalid pass-fail categorization. In practice, most grading systems appear to be multiple ($n > 2$) categorization scales which should do an excellent job on the important pass-fail distinction and a fair-to-poor job in linear prediction. If both of these objectives are important it would seem that something might be gained by a shift to a stanine scale (with no pretense that a given category indicated anything other than a rough indication of the <u>S</u>'s performance-ranking in a particular class). For each class the instructor could assign a pass-fail point based on the performance of that particular class. Thus, while every class would have the same distribution on the stanine scale, the percentage passing could range from zero to 100%. Little loss in flexibility of assignment of the pass-fail point occurs with such a system and prediction is improved with little added effort.

Appendix*

Suppose we wish to examine the degree to which grades given in the first semester of a 2-semester course correlate with grades given in the second semester of the course. In addition we wish to determine what effect expressing the grades in terms of stanine scores might have on such a validity coefficient.

Four pairs of classes were examined. In each case a raw score based on the sum of scores on all tests given during the semester was available. A stanine transformation was independently performed on these raw scores for each of the eight classes. The correlation between raw scores, between grades, and between stanine scores was then determined for those students who had attended both semesters. Results of these comparisons are presented in Table A1.

Table A1

Empirical Comparison of Validity Coefficients for Three Grading Scales

Course Level	Grades Used	N	Scale type: raw score	grade	stanine
graduate	A-B-C	34	.71	.57	.70
graduate	A-B-C	64	.59	.66	.72
graduate	A-B	18	.33	.01	.24
undergrad	A-B-C-D-E	21	.72	.78	.76

The data represent many practical problems met in real life. Most evident are effects of limitation of range due to selection as seen in the graduate course in which no grade lower than "B" was assigned. Even in the undergraduate course only 21 of the 71 students who took the first semester course also took the second semester. As might be expected these last students had noticeably higher grades (first semester median stanine score was four) than the average first semester student. In addition to the attenuating effect of reduced range, differentially skewed data produced nonlinearity of regression in several instances for either grades or raw scores.

While stanine scores did not show uniformly highest validity coefficients for these data they did account for an average of about 6% more variance than grades did. This improvement over grades is probably due to the joint effect of increased numbers of categories together with reduction in nonlinearity of regression. A linear transformation of raw scores to a 9-category scale would not be expected to have the latter benefits. An additional drawback of a simple linear transformation is that it requires more work than the stanine transformation.

*Empirical data were generously provided by Professors E. S. Avner, W. H. Batchelder, and W. E. Kappauf of the University of Illinois.

REFERENCES

Cronbach, L. J. & Gleser, G. C. Psychological tests and personnel decisions. Urbana: University of Illinois Press, 1957.

Ebel, R. L. The frequency of errors in the classification of individuals on the basis of fallible test scores. Educational and Psychological Measurement, 1947, (Winter), 725-734.

Ebel, R. L. Measuring educational achievement. Englewood Cliffs, N. J.: Prentice-Hall, 1965.

Kelly, T. L. Interpretation of educational measurements. Yonkers, N. Y.: World Book, 1927.

Kelly, T. L. Fundamentals of statistics. Cambridge, Mass.: Harvard University Press, 1947.

Pearson, K. Tables for statisticians and biometricians, part II. Cambridge: Cambridge University Press, 1931.

Peters, C. C. & Van Voorhis, W. R. Statistical procedures and their mathematical bases. New York: McGraw-Hill, 1940.

Willingham, W. W. An empirical evaluation of three types of college grade scales. Journal of Educational Measurement, 1964, 1(1), 69-72.

NAVY

3 Chief of Naval Research
Code 458
Department of the Navy
Washington, D.C. 20360

1 Director
ONR Branch Office
495 Summer Street
Boston, Massachusetts 02210

1 Director
ONR Branch Office
219 South Dearborn Street
Chicago, Illinois 60604

1 Director
ONR Branch Office
1030 East Green Street
Pasadena, California 91101

1 Contract Administrator
Southeastern Area
Office of Naval Research
2110 G Street, N.W.
Washington, D.C. 20037

10 Commanding Officer
Office of Naval Research
Box 39
Fleet Post Office
New York, New York 09510

1 Office of Naval Research
Area Office
207 West Summer Street
New York, New York 10011

1 Office of Naval Research
Area Office
1076 Mission Street
San Francisco, California 94103

6 Director
Naval Research Laboratory
Washington, D.C. 20390
Attn: Technical Information Division

20 Defense Documentation Center
Cameron Station, Building 5
5010 Duke Street
Alexandria, Virginia 22314

1 Superintendent
Naval Postgraduate School
Monterey, California 93940
Attn: Code 2124

1 Head, Psychology Branch
Neuropsychiatric Service
U. S. Naval Hospital
Oakland, California 94627

1 Commanding Officer
Service School Command
U. S. Naval Training Center
San Diego, California 92133

3 Commanding Officer
Naval Personnel Research Activity
San Diego, California 92152

1 Commanding Officer
Naval Air Technical Training Center
Jacksonville, Florida 32213

1 Officer in Charge
Naval Medical Neuropsychiatric
Research Unit
San Diego, California 92152

1 Dr. James J. Regan
Naval Training Device Center
Orlando, Florida 32813

1 Chief, Aviation Psychology Division
Naval Aerospace Medical Institute
Naval Aerospace Medical Center
Pensacola, Florida 32512

1 Chief, Naval Air Reserve Training
Naval Air Station
Box 1
Glenview, Illinois 60026

1 Chairman
Leadership/Management Committee
Naval Sciences Department
U. S. Naval Academy
Annapolis, Maryland 21402

1 Technical Services Division
National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 20014

1 Behavioral Sciences Department
Naval Medical Research Institute
National Naval Medical Center
Bethesda, Maryland 20014
Attn: Dr. W.W. Haythorn, Director

1 Commanding Officer
Naval Medical Field Research Laboratory
Camp Lejeune, North Carolina 28542

1 Director
Aerospace Crew Equipment Department
Naval Air Development Center, Johnsville
Warminster, Pennsylvania 18974

1 Chief, Naval Air Technical Training
Naval Air Station
Memphis, Tennessee 38115

1 Commander
Operational Test and Evaluation Force
U.S. Naval Base
Norfolk, Virginia 23511

1 Office of Civilian Manpower Management
Department of the Navy
Washington, D.C. 20350
Attn: Code 023

1 Chief of Naval Operations, Op-37
Fleet Readiness & Training Division
Department of the Navy
Washington, D.C. 20350

1 Chief of Naval Operations, Op-07TL
Department of the Navy
Washington, D.C. 20350

1 Capt. J.E. Rasmussen, MSC, USN
Chief of Naval Material (MAT 031M)
Room 1323, Main Navy Building
Washington, D.C. 20360

1 Naval Ship Systems Command, Code 03H
Department of the Navy
Main Navy Building
Washington, D.C. 20360

1 Chief
Bureau of Medicine and Surgery
Code 513
Washington, D.C. 20360

9 Technical Library
Bureau of Naval Personnel (Pers-11b)
Department of the Navy
Washington, D.C. 20370

3 Director
Personnel Research Laboratory
Washington Navy Yard, Building 200
Washington, D.C. 20390
Attn: Library

1 Commander, Naval Air Systems Command
Navy Department AIR-4133
Washington, D.C. 20360

1 Commandant of the Marine Corps
Headquarters, U. S. Marine Corps
Code A01B
Washington, D.C. 20380

ARMY

1 Human Resources Research Office
Division #6, Aviation
Post Office Box 428
Fort Rucker, Alabama 36360

1 Human Resources Research Office
Division #3, Recruit Training
Post Office Box 5787
Presidio of Monterey, California 93940
Attn: Library

1 Human Resources Research Office
Division #4, Infantry
Post Office Box 2086
Fort Benning, Georgia 31905

1 Department of the Army
U.S. Army Adjutant General School
Fort Benjamin Harrison, Indiana 46216
Attn: AGCS-EA

1 Director of Research
U.S. Army Armor Human Research Unit
Fort Knox, Kentucky 40121
Attn: Library

1 Dr. George S. Harker
Director, Experimental Psychology Division
U.S. Army Medical Research Laboratory
Fort Knox, Kentucky 40121

1 Research Analysis Corporation
McLean, Virginia 22101
Attn: Library

1 Human Resources Research Office
Division #5, Air Defense
Post Office Box 6021
Fort Bliss, Texas 79916

1 Human Resources Research Office
Division #1, Systems Operations
300 North Washington Street
Alexandria, Virginia 22314

1 Director
Human Resources Research Office
The George Washington University
300 North Washington Street
Alexandria, Virginia 22314

1 Armed Forces Staff College
Norfolk, Virginia 23511
Attn: Library

1 Chief
Training and Development Division
Office of Civilian Personnel
Department of the Army
Washington, D.C. 20310

1 U. S. Army Behavioral Science
Research Laboratory
Washington, D.C. 20315

1 Walter Reed Army Institute of
Research
Walter Reed Army Medical Center
Washington, D. C. 20012

1 Behavioral Sciences Division
Office of Chief of Research and
Development
Department of the Army
Washington, D.C. 20310

1 Dr. Vincent Cieri
U. S. Army Signal School
CAI Project
Fort Monmouth, New Jersey

AIR FORCE

1 Director
Air University Library
Maxwell Air Force Base
Alabama 36112
Attn: AUL-8110

1 Cadet Registrar (CRE)
U. S. Air Force Academy
Colorado 80840

1 Headquarters, ESD
ESVPT
L.G. Hanscom Field
Bedford, Massachusetts 01731
Attn: Dr. Mayer

1 6570 AMRL (MRHT)
Wright-Patterson Air Force Base
Ohio 45433
Attn: Dr. G. A. Eckstrand

1 Commandant
U.S. Air Force School of Aerospace
Medicine
Brooks Air Force Base, Texas 78235
Attn: Aeromedical Library
(SMSDL)

1 6570th Personnel Research
Laboratory
Aerospace Medical Division
Lackland Air Force Base
San Antonio, Texas 78236

1 AFOSR (SRLB)
1400 Wilson Boulevard
Arlington, Virginia 22209

1 Headquarters, U.S. Air Force
Chief, Analysis Division (AFPDPL)
Washington, D.C. 20330

1 Headquarters, U.S. Air Force
Washington, D. C. 20330
Attn: AFPTRTB

1 Headquarters, U.S. Air Force
AFRDDG
Room 1D373, The Pentagon
Washington, D.C. 20330

1 Research Psychologist
SCBB, Headquarters
Air Force Systems Command
Andrews Air Force Base
Washington, D.C. 20331

MISCELLANEOUS

1 Mr. Joseph J. Cowan
Chief, Personnel Research Branch
U.S. Coast Guard Headquarters
PO-1, Station 3-12
1300 E Street, N.W.
Washington, D.C. 20226

1 Director
Defense Atomic Support Agency
Washington, D.C. 20305

1 Executive Officer
American Psychological Association
1200 Seventeenth Street, N.W.
Washington, D.C. 20036

1 Dr. W. A. Bousfield
Department of Psychology
University of Connecticut
Stoors, Connecticut 06268

1 Dr. Lee J. Cronbach
School of Education
Stanford University
Stanford, California 94305

1 Professor L. E. Davis
Graduate School of Business
Administration
University of California, Los Angeles
Los Angeles, California 90024

1 Dr. Philip H. DuBois
Department of Psychology
Washington University
Lindell & Skinker Boulevards
St. Louis, Missouri 63130

1 Dr. Jack W. Dunlap
Dunlap and Associates
Darien, Connecticut 06820

1 Professor W. K. Estes
The Rockefeller University
New York, New York 10021

1 Dr. John C. Flanagan
American Institutes for Research
Post Office Box 1113
Palo Alto, California 94302

1 Dr. Frank Friedlander
Division of Organizational Sciences
Case Institute of Technology
Cleveland, Ohio 10900

1 Dr. Robert Glaser
Learning Research and Development
Center
Univeristy of Pittsburgh
Pittsburgh, Pennsylvania 15213

1 Dr. Bert Green
Department of Psychology
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

1 Dr. J. P. Guilford
University of Southern California
3551 University Avenue
Los Angeles, California 90007

1 Dr. Harold Gulliksen
Department of Psychology
Princeton University
Princeton, New Jersey 0540

1 Dr. M. D. Havron
Human Sciences Research, Inc.
Westgate Industrial Park
7710 Old Springhouse Road
McLean, Virginia 22101

1 Dr. Albert E. Hickey
Entelek, Incorporated
42 Plesant Street
Newburyport, Massachusetts 01950

1 Dr. William A. Hunt
Department of Psychology
Loyola University, Chicago
6525 North Sheridan Road
Chicago, Illinois 60626

1 Dr. Howard H. Kendler
Department of Psychology
University of California
Santa Barbara, California 93106

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
6780 Cortona Drive
Santa Barbara Research Park
Goleta, California 93107

1 Dr. A. B. Nadel
General Learning Corporation
5454 Wisconsin Avenue, N.W.
Washington, D.C. 20015

1 Dr. Slater E. Newman
Department of Psychology
North Carolina State University
Raleigh, North Carolina 27607

1 Dr. C. E. Noble
Department of Psychology
University of Georgia
Athens, Georgia 30601

1 Dr. Henry S. Odbert
National Science Foundation
1800 G Street, N.W.
Washington, D.C. 20550

1 Dr. Harry J. Older
Software Systems, Inc.
5810 Seminary Road
Falls Church, Virginia 22041

1 Dr. Leo J. Postman
Institute of Human Learning
University of California
2241 College Avenue
Berkeley, California 94720

1 Dr. Joseph W. Rigney
Electronics Personnel Research Group
University of Southern California
University Park
Los Angeles, California 90007

1 Dr. Arthur I. Siegel
Applied Psychological Services
Science Center
404 East Lancaster Avenue
Wayne, Pennsylvania 19087

1 Dr. Arthur W. Staats
Department of Psychology
Unversity of Hawaii
Honolulu, Hawaii 96822

1 Dr. Lawrence M. Stolurow
Harvard Computing Center
6 Appian Way
Cambridge, Massachusetts 02138

1 Dr. Donald W. Taylor
Department of Psychology
Yale University
333 Cedar Street
New Haven, Connecticut 06510

1 Dr. Ledyard R. Tucker
Department of Psychology
University of Illinois
Urbana, Illinois 61801

1 Dr. Karl L. Zinn
Center for Research on Learning
and Training
University of Michigan
Ann Arbor, Michigan 48104

1 Dr. James J. Asher
Department of Psychology
San Jose State College
San Jose, California 95114

1 Dr. Albert E. Goss
Department of Psychology
Douglass College, Rutgers
The State University
New Brunswick, New Jersey 08903

1 Mr. Halim Ozkaptan, Chief
Human Factors
Martin Company
Orlando, Florida 32809

1 Dr. Alvin E. Goins, Executive Secretary
Personality and Cognition Research
Review Committee
Behavioral Sciences Research Branch
National Institute of Mental Health
5454 Wisconsin Avenue, Room 10A11
Chevy Chase, Maryland 20203

1 Headquarters USAF (AFPTRD)
Training Devices and Instructional
Technology Division
Washington, D.C. 20330

1 Director
Education and Training Sciences
Department
Naval Medical Research Institute
Building 142
National Naval Medical Center
Bethesda, Maryland 20014

1 Dr. Mats Bjorkman
University of Umea
Department of Psychology
Umea 6, Sweden

1 LCDR J.C. Meredith, USN (Ret.)
Institute of Library Research
University of California, Berkeley
Berkeley, California 94720

1 Executive Secretariat
Interagency Committee on Manpower
Research
Room 515
1738 M Street, N.W.
Washington, D.C. 20036
Attn: Mrs. Ruth Relyea)

1 Dr. Marshall J. Farr
Assistant Director, Engineering
Psychology Program
Office of Naval Research (Code 455)
Washington, D.C. 20360

1 Mr. Joseph B. Blankenheim
NAVELEX 0474
Munitions Building, Rm. 3721
Washington, D.C. 20360

1 Technical Information Exchange
Center for Computer Sciences
and Technology
National Bureau of Standards
Washington, D.C. 20234

1 Technical Library
U. S. Naval Weapons Laboratory
Dahlgren, Virginia 22448

1 Technical Library
Naval Training Device Center
Orlando, Florida 32813

1 Technical Library
Naval Ship Systems Command
Main Navy Building, Rm. 1532
Washington, D.C. 20360

1 Technical Library
Naval Ordnance Station
Indian Head, Maryland 20640

1 Naval Ship Engineering Center
Philadelphia Division
Technical Library
Philadelphia, Pennsylvania 19112

1 Library, Code 0212
Naval Postgraduate School
Monterey, California 93940

1 Technical Reference Library
Naval Medical Research Institute
National Naval Medical Center
Bethesda, Maryland 20014

1 Technical Library
Naval Ordnance Station
Louisville, Kentucky 40214

1 Library
Naval Electronics Laboratory Center
San Diego, California 92152

1 Technical Library
Naval Undersea Warfare Center
3202 E. Foothill Boulevard
Pasadena, California 91107

1 Dr. Russ L. Morgan (HRHT)
Training Research Division
Human Resources Laboratory
Wright-Patterson Air Force Base
Ohio 45433

1 Headquarters, Air Training Command
Randolph Air Force Base, Texas
78148
Attn: ATXTD (Dr. Meyer)

1 Mr. Michael Macdonald-Ross
International Training and Education
Company Limited
ITEC House
29-30 Ely Place
London EC1
ENGLAND

1 Commanding Officer
U. S. Naval Schools Command
Mare Island
Vallejo, California 94592

1 Dr. Don C. Coombs, Assistant Director
ERIC Clearinghouse
Stanford University
Palo Alto, California 94305

1 CDR H. J. Connery, USN
Scientific Advisory Team (Code 71)
Staff, COMASWFORLANT
Norfolk, Virginia 23511

1 ERIC Clearinghouse
Educational Media and Technology
Stanford University
Stanford, California

1 ERIC Clearinghouse
Vocational and Technical Education
Ohio State University
Columbus, Ohio 43212

1 Dr. Benton J. Underwood
Department of Psychology
Northwestern University
Evanston, Illinois 60201

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Illinois, Board of Trustees Computer-based Education Research Laboratory Urbana, Illinois	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

OBJECTIVE CRITERIA FOR EVALUATION OF GRADING SCALES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

R. A. Avner

6. REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO OF REFS
August 1970	14	8

8a. CONTRACT OR GRANT NO ONR Nonr 3985(08) b. PROJECT NO c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) CERL Report X-18 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Distribution of this report is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency, Office of Naval Research

13. ABSTRACT

Maximum linear prediction, maximum total correct classifications for a group, and maximum probability of correct classification for an individual are compared as objective criteria for univariate grading scales. The goals of valid prediction and valid classification are seen as leading to conflicting criteria. A compromise measure consisting of a standardized score with variable pass-fail point is suggested.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Security Classification

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Measurement						
Evaluation						
Education						
Training						
Univariate Measures						
Classification						
Computer-based education						
Computer-assisted instruction						
PLATO						

DD FORM 1 NOV 65 1473 (BACK)

Security Classification